A PRACTICAL ANALYTIC METHOD FOR CALCULATING $\pi(x)$

JENS FRANKE, THORSTEN KLEINJUNG, JAN BÜTHE, AND ALEXANDER JOST

Abstract. In this paper we give a description of a practical analytic method for the computation of $\pi(x)$, the number of prime numbers $\leq x$. The method is similar to the one proposed by Lagarias and Odlyzko but uses the Weil-Barner explicit formula instead of curve integrals.

So far, practical calculations of the prime counting function $\pi(x)$ seem to have been based upon combinatorial methods going back to [Mei70] and refined by [Leh59] and [LMO85]. The fast evaluation of multiple values of the Riemann zeta function made it possible in theory to describe an analytic method faster than the combinatorial methods [LO87]. This was further analyzed by Galway [Gal04], however without implementing it.

In this paper, we describe a version of the analytic method which uses the Weil explicit formula. We implemented this method and, assuming the Riemann hypothesis, obtained the following values of $\pi(x)$, which, as far as we know, were unknown before:

<table>
<thead>
<tr>
<th>$x$</th>
<th>Analytic approximation to $\pi(x)$ on RH</th>
<th>$\varepsilon$</th>
<th>$c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^{76}$</td>
<td>1462626667154509638735-6.60903e-09</td>
<td>3e-10</td>
<td>30</td>
</tr>
<tr>
<td>$2^{77}$</td>
<td>2886507381056867953916-1.72698e-08</td>
<td>3e-10</td>
<td>30</td>
</tr>
<tr>
<td>$2^{78}$</td>
<td>5697549648954257752872-2.98507e-09</td>
<td>3.3e-10</td>
<td>30</td>
</tr>
<tr>
<td>$2^{79}$</td>
<td>11248065615133675809379+9.5313e-09</td>
<td>3.3e-10</td>
<td>33</td>
</tr>
<tr>
<td>$1e24$</td>
<td>1843559767349200867866+3.3823e-08</td>
<td>3e-10</td>
<td>30</td>
</tr>
<tr>
<td>$2^{80}$</td>
<td>22209558889635384205844+2.10808e-09</td>
<td>3.3e-10</td>
<td>33</td>
</tr>
</tbody>
</table>

The method treated in Theorem 2.7 has been used, and the last two columns of the table give the values of the parameters $\varepsilon$ and $c$ used in the calculations. The results depend on the Riemann hypothesis; it is thus necessary to apply Theorem 2.7 with $h = .5$ in order to obtain a bound on the remainder which is $< .5$. Meanwhile the value $\pi(10^{24})$ has been confirmed without the assumption of the Riemann hypothesis by a calculation of D. Platt [Pla15], using an analytic method based on the aforementioned work of Galway.

Whether or not the correctness of the result of a calculation of $\pi(x)$ using one of the methods described in this paper assumes the Riemann hypothesis depends on which of the two methods is being used and also on the choice of parameters. With choices of parameters not assuming the Riemann hypothesis, analytic calculations of $\pi(x)$ have been done for values as large as $x = 10^{20}$.

Received by the editor October 7, 2013 and, in revised form, November 11, 2014.
2010 Mathematics Subject Classification. Primary 11Y35; Secondary 11Y70.

©2017 American Mathematical Society
For our calculations, we determined all zeros of $\zeta\left(1/2 + it\right)$ with $|t| < 10^{11}$ to an absolute precision of $2^{-64}$. This was done on various CPUs provided by the Hausdorff Center for Mathematics and the Institute for Numeric Simulation at Bonn University and took an estimated 130 CPU days on these machines. To obtain more precise information, part of the zero-finding jobs were rerun on a single 2.66 GHz CPU and the CPU time extrapolated, giving an estimated 3,297 hours of CPU time on that machine for the calculation of all zeros with imaginary parts between $2\pi 10^6$ and $10^{11}$. Calculation of the zeros with smaller positive imaginary parts takes an insignificant amount of time. The computation of values of $\pi(x)$ for individual $x$ then requires additional calculations, which in the case of $10^{24}$ take about 1,600 CPU hours on that CPU. This is explained in more detail at the end of subsection 4.3.

1. DESCRIPTION OF THE METHOD

Let $\tilde{\pi}(x)$ be the number of primes $< x$, and let $\pi(x)$ be equal to $\tilde{\pi}(x)$ if $x$ is not a prime number and to $1/2 + \tilde{\pi}(x)$ otherwise. Let $Ei(z)$ be the antiderivative of $e^z/z$ which is holomorphic outside $(-\infty, 0)$ and which equals

$$\lim_{\varepsilon \downarrow 0} \left( \int_{-\infty}^{-\varepsilon} + \int_{\varepsilon}^{z} \right) \frac{e^t}{t}$$

for positive real $z$. Recall the well-known (or see Lemma 2.6) asymptotic expansion

$$Ei(z) \sim \pi \cdot i \text{ sgn}(\text{Im } z) + \sum_{k=1}^{\infty} \frac{(k-1)!}{z^k} e^z.$$ (1.1)

Our starting point is the classical explicit formula formulated by Riemann [Rie59] and proved by von Mangoldt [vM95]:

$$\pi^*(x) = \text{li}(x) - \sum_{\rho} \pi(\rho \log(x)) - \log(2) + \int_{x}^{\infty} \frac{dt}{t(t^2 - 1) \log t},$$ (1.2)

where $\pi^*$ is the Riemann prime counting function

$$\pi^*(x) = \sum_{k=1}^{\infty} \frac{\pi(\sqrt[k]{x})}{k}$$

and $\sum_{\rho} f(\rho)$ is a sum over zeros of the zeta-function in the strip $0 < \text{Re}(s) < 1$ calculated as

$$\lim_{T \rightarrow \infty} \sum_{|\text{Im } \rho| < T} n_{\rho} f(\rho),$$

$n_{\rho}$ being the multiplicity of the zero $\rho$.

For a derivation of (1.2) from the Weil-Barner explicit formula, see [BFJK13]. It is well known that (1.2) converges too slowly to be of direct practical use for calculating $\pi(x)$ [RG70]. To obtain a practical method, we subtract a sum over zeros of a function with a similar asymptotic behaviour as the summand in (1.2), but which can be transformed by the Weil-Barner formula into a sum over prime powers in an interval which is much shorter than $[0, x]$. Recall the Weil-Barner formula

$$w_s(\hat{g}) = w_f(g) + w_{-\infty}(g),$$ (1.3)
where we use the convention
\begin{equation}
F g(t) = \hat{g}(t) = \int_{-\infty}^{\infty} e^{i\pi t} g(x) \, dx,
\end{equation}
\[
g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\pi t} \hat{g}(t) \, dt
\]
for the Fourier transform \( \hat{g} \) of \( g \), and put
\[
(1.5) \quad w_s(\hat{g}) = -\hat{g}(\frac{i}{2}) - \hat{g}(-\frac{i}{2}) + \sum_{\rho} \hat{g}(\frac{\alpha}{\pi} - \frac{1}{2}) ,
\]
\[
w_f(g) = -\sum_{p} \sum_{m=1}^{\infty} \frac{\log(p)}{p^{m/2}} \left( g(m \log p) + g(-m \log p) \right) ,
\]
\[
w(\gamma) = -\int_{0}^{\infty} \left( g(t) + g(-t) - 2g(0) \right) e^{-t/2} dt \left( \frac{\Gamma'}{\Gamma} (\frac{1}{4} - \log(\pi)) g(0) .
\]

The notation is motivated by the well-known analogy with the Selberg-Arthur trace formula where the sum over zeros \( w_s \) is similar to the spectral side, \( w_f \) is the contribution of the finite primes and \( w_\infty \) the contribution of the “infinite prime”.

This was shown to hold by Barner [Bar81], motivated by work of Weil [Web52] and others, for \( \mathbb{C} \)-valued functions \( g \) on \( \mathbb{R} \) such that:

- For some \( \alpha > 0 \), the function \( g(t)e^{(\frac{1}{2} + \alpha)|t|} \) has bounded variation on \( \mathbb{R} \).
- At its locations of discontinuity, \( g \) is normalised in the sense that \( g(t) = \frac{1}{2}(g(t^+) + g(t^-)) \), with \( g(t^+) = \lim_{\varepsilon \downarrow 0} g(t + \varepsilon) \),
- For some positive \( \varepsilon \), we have \( 2g(0) = g(t) + g(-t) + O(|t|^{\varepsilon}) \) as \( t \to 0 \).

We call the set of these functions the Barner class throughout this paper.

To construct a function \( \Phi \) with compact support such that the summands of \( w_s(\Phi) \) cancel the main asymptotic term of the sum over zeros in (1.2), let
\[
l(\xi) = \frac{1}{\sinh c} \frac{\sin(\sqrt{\xi^2 - c^2})}{\sqrt{\xi^2 - c^2}}.
\]

This function has been proposed by Logan [Log88] to minimize \( \int_{|\xi|>c} \frac{l(\xi)}{|\xi|} d\xi \) among all \( l \) with \( \int_{-\infty}^{\infty} l(\xi) d\xi = 1 \) and with Fourier image supported in \([-1,1]\). Let \( \Phi_{x,c,\varepsilon} \) be given by
\begin{equation}
(1.6) \quad \tilde{\Phi}_{x,c,\varepsilon}(\xi) = \frac{x^{\frac{1}{2} + i\xi}}{\log x} \left( -\frac{i}{\xi} + \frac{1}{\xi^2} \left( \frac{1}{2} - \frac{1}{\log x} \right) \right) (l(\varepsilon \xi) - 1) .
\end{equation}

The last factor has a second order zero at \( \xi = 0 \), cancelling the poles of the first factor there. Using the Paley-Wiener Theorem, it is easy to see that the support of \( \Phi_{x,c,\varepsilon} \) is contained in \([\log x - \varepsilon, \log x + \varepsilon]\). When \( \xi \) is a real number > \( \frac{c}{\varepsilon} \), then \( |l(\varepsilon \xi)| < \frac{c}{\sinh c} \). With values like \( c = 30 \) for \( x = 10^{24} \), this means that for \( \xi > \frac{c}{\varepsilon} \), (1.6) is sufficiently close to
\[
\frac{x^{\frac{1}{2} + i\xi}}{\log x} \left( \frac{i}{\xi} - \frac{1}{\xi^2} \left( \frac{1}{2} - \frac{1}{\log x} \right) \right) .
\]
for the purpose of calculating \( \pi(x) \), cancelling the first two terms,
\[
e^{(\frac{1}{2} + i\xi)} \log x \left( \frac{1}{(\frac{1}{2} + i\xi) \log x} + \frac{1}{(\frac{1}{2} + i\xi)^2 (\log x)^2} \right)
\]
\[
= e^{(\frac{1}{2} + i\xi)} \log x \left( -\frac{i}{\xi \log x} + \left( \frac{1}{2\xi^2 \log x} - \frac{1}{\xi^2 (\log x)^2} \right) + O(\frac{1}{\xi^2}) \right),
\]
of the asymptotic expansion (1.1) for the summand \( \text{Ei} \left( (\frac{1}{2} + i\xi) \log x \right) \) of (1.2). A
detailed estimate of the remainder will be given in the next section.

Similarly, one could take
\[
\hat{\Psi}_{x,c,\epsilon,\delta}(\xi) = \frac{x}{\log x} \left( -\frac{i}{\xi} + \frac{1}{\xi^2} \left( \frac{1}{2} - \frac{1}{\log x} \right) \right.
\]
\[
\left. + \frac{i}{\xi^3} \left( \frac{1}{4} - \frac{1}{\log x} + \frac{2}{(\log x)^2} \right) \right) (g_{c,\delta}(\xi) - 1),
\]
essentially cancelling the first three summands of (1.2), where
\[
g_{c,\delta}(\xi) = l_c(\xi) (a_\delta + b_\delta \cos(\delta \xi))
\]
with \( b_\delta = -l''_c(0) / \delta^2 \) and \( a_\delta = 1 - b_\delta \).

The method is thus to calculate \( \pi^*(x) \) as the sum of \( w_f(f) \) and
\[
\text{li}(x) + \hat{f}(i/2) + \hat{f}(-i/2) - \sum_{\rho} (\text{Ei}(\rho \log x) + \hat{f}(\rho i - \frac{1}{2i}))
\]
\[
- \log(2) + \int_x^\infty \frac{dt}{t(t^2 - 1) \log t} + w_\infty(f),
\]
where \( f \) is either \( \Phi_{x,c,\epsilon} \) or \( \Psi_{x,c,\epsilon,\delta} \). The infinite sum over zeros in (1.8) is truncated
to all zeros with imaginary part < \( \frac{c}{\epsilon} \) and the term \( w_\infty(f) \) omitted:
\[
\text{li}(x) + \hat{f}(i/2) + \hat{f}(-i/2) - \sum_{|\text{Im}(\rho)| < T} (\text{Ei}(\rho \log x) + \hat{f}(\rho i - \frac{1}{2i}))
\]
\[
- \log(2) + \int_x^\infty \frac{dt}{t(t^2 - 1) \log t}.
\]

The functions \( \Phi_{x,c,\epsilon} \) and \( \Psi_{x,c,\epsilon,\delta} \) have been proposed by the first author.

2. Estimation of the remainder

2.1. The distribution of zeta-zeros. In order to estimate the remainder
\[
\sum_{|\text{Im}(\rho)| > T} (\text{Ei}(\rho \log x) + \hat{f}(\rho i - \frac{1}{2i}))
\]
some information on the number of zeros of the Riemann zeta function is needed.
Let \( N(T) \) be the number, counted by multiplicity, of zeros \( \rho \) of the zeta function
satisfying \( 0 < \text{Im}(\rho) \leq T \). The asymptotic behaviour of \( N(T) \) is well known. We
use Turing’s \( \Theta \)-notation and say \( g(t) = \Theta(f(t)) \) if \( |g(t)| \leq f(t) \).

Theorem 2.1 (Rosser). Let \( T \geq 2 \). Then we have
\[
N(T) = \frac{T}{2\pi} \log \frac{T}{2\pi} - \frac{T}{2\pi} + \frac{7}{8} + \Theta(0.137 \log T + 0.443 \log \log T + 1.588).
\]

Proof. See [Ros41] p. 223. □
This result is more accurate than necessary, so the following two corollaries will be used instead.

**Corollary 2.2.** Let \( T > 10^6 \). Then we have

\[
N(T) = \frac{T}{2\pi} \log \frac{T}{2\pi e} + \Theta(0.4 \log T).
\]

**Proof.** Obvious. \( \square \)

**Corollary 2.3.** Let \( T > 10^6 \) and \( M > 0 \). Then we have

\[
N(T + M) - N(T) \leq \frac{M + 6}{2\pi} \log \frac{T + M}{2\pi}.
\]

**Proof.** Since \( 0.137 \log T + 0.443 \log \log T + 1.588 = \Theta(0.35 \log T) \) for \( T > 10^6 \) and since \( \log \left( \frac{T + M}{T} \right) < \frac{M}{T} \), equation (2.1) implies

\[
N(T + M) - N(T) = \frac{M + 6}{2\pi} \log \left( \frac{T + M}{2\pi} \right).
\]

2.2. Auxiliary lemmas.

**Lemma 2.4.** Let \( 0 < \varepsilon < 10^{-5} \) and \( c \geq 10 \). Then we have

\[
\sum_{\rho \in \rho} \frac{|l_c(\varphi (\xi - \frac{1}{2}))|}{|\operatorname{Im}(\rho)|} \leq 0.65 e^{c(3/4 - 1)} \log(3c) \log(c/\varepsilon).
\]

**Proof.** First we prove the following bound for Logan’s function. Let \( |x| \geq c \) and \( |y| \leq 10^{-5} \). Then, we have

\[
|l_c(x + iy)| \leq \frac{c}{\sinh(c)} e^{c(3/4)} \min \left\{ 1, \frac{1}{|x| - c} \right\}.
\]

Since \( l_c \) is even and interchanges with complex conjugation, we may restrict ourselves to the case where \( x \) and \( y \) are non-negative. Let \( z = x + iy \) and

\[
f(z) = e^{-c \frac{\sqrt{y^2 - c^2}}{4}} \frac{\sin(\sqrt{y^2 - c^2})}{\sqrt{y^2 - c^2}}.
\]

First we show that the function \( x \mapsto \operatorname{Im}(\sqrt{x^2 - c^2}) \) is monotonously decreasing for \( x > 0 \) and \( y \geq 0 \) if we take the main branch of the logarithm to define the square root. For \( y = 0 \) this is trivial, so we may assume \( y > 0 \). Now let \( (a + ib)^2 = z^2 - c^2 \).

Then, since \( b > 0 \), one easily obtains the equality

\[
x^2 y^2 - b^2 = x^2 - y^2 - c^2.
\]

So \( b \) tends to \( \sqrt{y^2 + c^2} > y \) if \( x \) tends to 0, and it tends to \( y \) if \( x \) tends to \( \infty \). Furthermore \( b \), as a function of \( x \), has no local minimum in \((0, \infty)\), since this would imply \( b^2 = y^2 \), which contradicts (2.5). Consequently, we have

\[
0 \leq \operatorname{Im}(\sqrt{z^2 - c^2}) \leq \operatorname{Im}(\sqrt{(c + iy)^2 - c^2}) = \sqrt{2c y - y^2} \sin \left( \frac{\pi}{4} + \frac{1}{2} \arctan \left( \frac{y}{2c} \right) \right)
\]

\[
\leq \sqrt{2c + 10^{-5}} y \sin \left( \frac{\pi}{4} + \frac{1}{4 \cdot 10^6} \right) \leq 0.71 \sqrt{2cy} \leq \frac{c \sqrt{2cy}}{4}.
\]
for $x \geq c$, which follows easily by monotonity arguments. So if we take into account that

$$\left| \frac{\sin(w)}{w} \right| = \left| \frac{1}{2} \int_{-1}^{1} e^{itw} dt \right| \leq e^{\text{Im}(w)},$$

we see that $|f(z)| \leq 1$ for $c \leq x \leq c + 1$.

For $x \geq c + 1$, we have $|\sin \sqrt{z^2 - c^2}| \leq e^{\frac{c}{2}|z|}$ by (2.6). Furthermore, since we assumed $c \geq 10$, we have

$$\left| \sqrt{z^2 - c^2} \right|^2 \geq |\text{Re}(z^2 - c^2)| = (x - c)^2 + 2c(x - c) - y^2 \geq (x - c)^2$$

and we arrive at

$$|f(z)| \leq \frac{1}{x - c} \quad \text{for} \quad x \geq c + 1,$$

which proves (2.4).

We will now estimate the sum

$$\sum_{\rho \in S} \frac{|f(\frac{\rho}{\varepsilon} - \frac{1}{2\pi})|}{|\text{Im}(\rho)|}.$$

By Corollary 2.3 there are at most

$$\frac{\varepsilon^{-1} + 6}{2\pi} \log \frac{c + 1}{2\pi \varepsilon}$$

zeros with imaginary part in $(\frac{c}{\varepsilon}, \frac{c + 1}{\varepsilon})$. Therefore, we have

$$\sum_{\xi < \text{Im}(\rho) \leq \frac{c + 1}{\varepsilon}} \frac{|f(\frac{\rho}{\varepsilon} - \frac{1}{2\pi})|}{|\text{Im}(\rho)|} \leq \frac{\varepsilon^{-1} + 6}{c} \frac{\log \frac{c + 1}{2\pi \varepsilon}}{2\pi} \leq \frac{0.16}{c} \log \frac{c}{\pi \varepsilon}.$$

For the remaining zeros we use the bound in (2.7) and get

$$\sum_{\text{Im}(\rho) > \frac{c + 1}{\varepsilon}} \frac{|f(\frac{\rho}{\varepsilon} - \frac{1}{2\pi})|}{|\text{Im}(\rho)|} \leq - \int_{\frac{c + 1}{\varepsilon}}^{\infty} N(t) \frac{d}{dt} \left( \frac{1}{t(\varepsilon t - c)} \right) dt - \frac{\varepsilon}{c + 1} N\left( \frac{c + 1}{\varepsilon} \right)$$

$$\leq - \int_{\frac{c + 1}{\varepsilon}}^{\infty} N(t) \frac{d}{dt} \left( \frac{1}{t(\varepsilon t - c)} \right) dt - \frac{1}{2\pi} \log \frac{c + 1}{2\pi \varepsilon}$$

$$+ \Theta \left( 0.4 \frac{\varepsilon}{c} \log \frac{c + 1}{\varepsilon} \right) \quad (2.9)$$

by partial summation, where we also used the estimate in Corollary 2.2 for $N(t)$. We use that same estimate to further evaluate the integral on the right hand side of (2.9). For the main term $\frac{1}{2\pi} \log \frac{t}{2\pi \varepsilon}$ therein we get

$$- \frac{1}{2\pi} \int_{\frac{c + 1}{\varepsilon}}^{\infty} t \log \frac{t}{2\pi \varepsilon} dt \frac{d}{dt} \left( \frac{1}{t(\varepsilon t - c)} \right) dt = \frac{1}{2\pi} \log \frac{c + 1}{2\pi \varepsilon} + \frac{1}{2\pi} \int_{\frac{c + 1}{\varepsilon}}^{\infty} \log \left( \frac{t}{2\pi \varepsilon} \right) dt.$$
Here the first term cancels the second term on the right hand side of (2.9). We continue estimating the integral
\[
\frac{1}{2\pi} \int_{\frac{c}{\sqrt{\pi}}}^{\infty} \frac{\log \frac{t}{\pi}}{t(\pi t - c)} dt = \frac{1}{2\pi} \int_{1}^{c} \frac{\log \left( \frac{t}{2\pi} \right)}{(t + c)t} dt
\]
\[
\leq \frac{1}{2\pi} \left( \frac{\log \left( \frac{c}{\pi} \right)}{c} \int_{1}^{c} \frac{dt}{t} + \int_{c}^{\infty} \frac{\log \left( \frac{t}{2\pi} \right)}{t^2} dt \right)
\]
\[
= \frac{1}{2\pi c} \log \left( \frac{c}{\pi} \right) \cdot \log c + \frac{\log \frac{2c}{2\pi}}{2\pi c} + \frac{1}{2\pi} \int_{c}^{\infty} \frac{dt}{t(t + c)}
\]
(2.10)

For the contribution of the Θ-term in (2.2) to the integral in (2.9) the same substitution shows that
\[
- \int_{\frac{c}{\sqrt{\pi}}}^{\infty} \log t \frac{d}{dt} \left( \frac{1}{t(\pi t - c)} \right) dt = \varepsilon \int_{1}^{\infty} \frac{\log \left( \frac{t + \varepsilon}{\varepsilon} \right)}{(t + c)^2 t} dt + \varepsilon \int_{1}^{\infty} \frac{\log \left( \frac{t + \varepsilon}{\varepsilon} \right)}{(t + c)t^2} dt
\]
\[
\leq 2\varepsilon \int_{1}^{\infty} \frac{\log \left( \frac{t + \varepsilon}{\varepsilon} \right)}{(t + c)^2 t^2} = 2\varepsilon \int_{1}^{\infty} \frac{\log(t + c) - \log \varepsilon}{(t + c)t^2}
\]
\[
\leq 2\varepsilon \int_{1}^{\infty} \frac{dt}{t^2} - 2\varepsilon \log \varepsilon \int_{1}^{\infty} \frac{dt}{t^3} = 2\varepsilon - \varepsilon \log \varepsilon \leq \varepsilon \log \frac{c}{\varepsilon}.
\]
Combining this with (2.8) and (2.9), where we generously bound the Θ-term by \(2\varepsilon \log(c/\varepsilon)\), and (2.10), we therefore arrive at
\[
\sum_{\rho \in \rho, \text{Im}(\rho) > \varepsilon} \left| \frac{f(\varepsilon \left( \frac{\rho}{\text{Im}(\rho)} \right))}{\text{Im}(\rho)} \right| \leq \frac{1}{c} \left[ 0.16 \log \frac{c}{\varepsilon} + 0.16 \log \left( \frac{c}{\varepsilon} \right) \log(c) + 3\varepsilon \log \left( \frac{c}{\varepsilon} \right) \right]
\]
\[
\leq 0.16 \log(3c) \log \frac{c}{\varepsilon}.
\]
From this and (2.3), the result follows if we take into account that \(\sinh(c)^{-1} \leq 2.001 \exp(-c)\) for \(c \geq 10\) and that the set of zeta zeros is invariant under complex conjugation. \(\square\)

**Lemma 2.5.** Let \(T > 10^6\) and \(k \geq 2\). Then we have
\[
(2.11) \sum_{\rho \in \rho, \text{Im}(\rho) > T} \frac{1}{|\text{Im}(\rho)|^k} \leq \frac{1}{k - 1} \frac{T^{1-k}}{\pi \log(T)} + \frac{T^{1-k}}{10000}.
\]

**Proof.** This follows easily from the identity
\[
(2.12) \sum_{\rho \in \rho, \text{Im}(\rho) > T} \frac{1}{|\text{Im}(\rho)|^k} = k \int_{T}^{\infty} \frac{N(t)}{t^{k+1}} dt - \frac{N(T)}{T^k}
\]
and Corollary 2.2, where we use
\[
k \int_{T}^{\infty} \frac{\log \left( \frac{t}{2\pi e} \right)}{t^k} dt = \frac{k}{k - 1} \cdot \frac{\log \left( \frac{T}{2\pi e} \right)}{T^{k-1}} + \frac{k}{(k - 1)^2} \cdot \frac{1}{T^{k-1}}
\]
for the main term of (2.2) and
\[
0.4 \cdot k \int_T^{\infty} \frac{\log t}{t^{k+1}} dt = \frac{0.4}{T^k} (\log(T) + \frac{1}{k})
\]
for the integral of the remainder of (2.2). □

**Lemma 2.6.** For \( z \in \mathbb{C} \setminus (-\infty, \infty) \) let \( R(k, z) \) be defined by the identity
\[
\text{Ei}(z) = \text{sgn}(\text{Im}(z))\pi i + \sum_{l=1}^{k} (l-1)! \frac{e^z}{z^l} + R(k, z).
\]
Then we have
\[
|R(k, z)| \leq \frac{2k!}{|\text{Im}(z)|^{k+1}} e^{\text{Re}(z)}.
\]

**Proof.** Let \( z = x + iy \). Since \( \text{Ei} \) interchanges with complex conjugation, we may assume \( y > 0 \). For such \( z \) let \( \gamma_+(z) \) be the polygonal chain defined by the sequence \((-\infty + iR, x + iR, z)\) and \( \gamma_-(z) \) the polygonal chain defined by the sequence \((-\infty - iR, 1 - iR, 1 + iy, z)\). Then
\[
\frac{1}{2} \left( \int_{\gamma_+(1)} + \int_{\gamma_-(1)} \right) e^\xi \frac{d\xi}{\xi} = \frac{1}{2} \lim_{R \to 0} \left( \int_{\gamma_+(1)} + \int_{\gamma_-(1)} \right) e^\xi \frac{d\xi}{\xi} = \int_{-\infty}^{1} \frac{e^t}{t} dt + \lim_{R \to 0} \int_{0}^{1} \frac{\sinh(t + iR)}{t + iR} + \frac{\sinh(t - iR)}{t - iR} dt = \int_{-\infty}^{1} \frac{e^t}{t} dt + 2 \int_{0}^{1} \frac{\sinh(t)}{t} dt = \text{Ei}(1)
\]
by the Dominated Convergence Theorem. Therefore, we have
\[
\text{Ei}(z) = \frac{1}{2} \left( \int_{\gamma_+(z)} e^\xi \frac{d\xi}{\xi} + \int_{\gamma_-(z)} e^\xi \frac{d\xi}{\xi} \right).
\]
By taking the limit \( R \to \infty \) we find
\[
\int_{\gamma_+(z)} e^\xi \frac{d\xi}{\xi} = - \int_{z}^{x+i\infty} e^\xi \frac{d\xi}{\xi}
\]
and
\[
\int_{\gamma_-(z)} e^\xi \frac{d\xi}{\xi} = \int_{C} e^\xi \frac{d\xi}{\xi} - \int_{x+i\infty}^{z} e^\xi \frac{d\xi}{\xi},
\]
where \( C \) is a path connecting \( x - i\infty \) and \( x + i\infty \) which intersects the real axis once at a positive value. By shifting \( C \) towards \(-\infty \) and collecting the residue at 0 we find that
\[
\int_{C} e^\xi \frac{d\xi}{\xi} = 2\pi i
\]
and therefore obtain the identity
\[
\text{Ei}(z) = \pi i - i \int_{y}^{\infty} \frac{e^{x+it}}{x+it} dt.
\]
By partial integration we find that
\[ -i \int_{y}^{\infty} \frac{e^{x+it}}{x+it} \, dt = \sum_{k=0}^{n} \frac{1}{(x+it)^{k+1}} \int_{y}^{\infty} \frac{e^{x+it}}{(x+it)^{k+1}} \, dt. \]
Consequently, we get
\[ R(k, z) = -ik! \int_{y}^{\infty} \frac{e^{x+it}}{(x+it)^{k+1}} \, dt, \]
and so we find
\[ |R(k, z)| \leq k! \frac{e^{x}}{|z|^{k+1}} + (k+1)!e^{x} \int_{y}^{\infty} \frac{1}{|x+it|^{k+2}} \, dt \leq \frac{2k!}{y^{k+1}} e^{x}. \]

2.3. Results.

**Theorem 2.7** (Method I). Let \( x \geq 10^{10}, \varepsilon \leq 10^{-5} \) and \( c \geq 10 \). Then we have
\[ \sum_{|\Im(\rho)| > \frac{\varepsilon}{2}} \left( \text{Ei}(\rho \log x) + \hat{\Phi}_{x,c,\varepsilon} \left( \frac{\rho}{\pi} - \frac{1}{2l} \right) \right) \]
\[ = \Theta \left( \frac{x^{h}}{\log x} \left( \frac{c}{\varepsilon} \right) \left( 0.05 \left( \frac{\varepsilon}{c} \right)^{2} + 0.66 \log(3c) e^{(\sqrt{\varepsilon/4}-1)c} \right) \right), \]
where \( h = \frac{1}{2} \) if the Riemann hypothesis is assumed and \( h = 1 \) otherwise.

**Proof.** We recall that
\[ \hat{\Phi}_{x,c,\varepsilon} = \frac{x^{1/2+i\varepsilon}}{\log x} \left( -i \frac{1}{\xi} + \frac{1}{\xi^{2}} \left( \frac{1}{2} - \frac{1}{\log x} \right) \right) (l_{c}(\varepsilon \xi) - 1). \]
A straightforward computation shows that
\[ \frac{1}{(\frac{1}{2} + i\xi)l} + \frac{1}{((\frac{1}{2} + i\xi)l)^{2}} - \frac{1}{l} \left( -i + \frac{1}{\xi^{2}} \left( \frac{1}{2} - \frac{1}{l} \right) \right) = \frac{-2 + l - 8i\xi + 2il\xi}{2l^{2}(l - 2\xi)^{2}\xi^{2}} \]
holds, where we abbreviated \( \log x \) by \( l \). Without loss of generality we can assume \( \text{Re}(\xi) > 0 \). Under the conditions imposed on \( x, \varepsilon \) and \( c \) we can assume \( l \geq 23, \text{Re}(\xi) \geq 10^{6} \), and since the non-trivial zeros of \( \zeta(s) \) have real part in \((0, 1)\) we can further assume \( |\Im(\xi)| < 1/2 \). Under these conditions, we have \( |\xi| \leq 1.0001 \text{Re}(\xi) \) and \( |l - 2\xi| \geq 2 \text{Re}(\xi) \). Consequently, we get
\[ \frac{2 + l + 8.001 \text{Re}(\xi) + 2.001l \text{Re}(\xi)}{8 \text{Re}(\xi)^{3}l^{2}} \]
\[ \leq \frac{1}{8 \text{Re}(\xi)^{3}l} \left( \frac{2/l + 1}{\text{Re}(\xi)} + \frac{8.001}{l} + 2.001 \right) \leq \frac{0.3}{l \text{Re}(\xi)^{3}}. \]
On the other hand, using the same bounds for \( \xi \) and \( l \), we get
\[ \text{Ei}((1/2 + i\xi)l) - \text{sgn}(\text{Re}(\xi)) \pi i \]
\[ = x^{1/2+i\xi} \left( \frac{1}{(1/2 + i\xi)l} + \frac{1}{((1/2 + i\xi)l)^{2}} + \Theta \left( \frac{0.01}{l \text{Re}(\xi)^{3}} \right) \right) \]
from Lemma 2.6. Also, we have
\[ \left| \frac{1}{\log x} \left( -i \frac{1}{\xi} + \frac{1}{\xi^{2}} \left( \frac{1}{2} - \frac{1}{\log x} \right) \right) l_{c}(\varepsilon \xi) \right| \leq 1.01 \frac{|l_{c}(\varepsilon \xi)|}{l |\text{Re}(\xi)|}. \]
Combining the last inequality with (2.15), (2.16) and (2.17) we obtain the inequality

\[
(\text{2.19}) \quad \left| \text{Ei}((1/2 + i\xi) \log x) - \text{sgn} (\text{Re}(\xi)) \pi i + \hat{\Phi}_{x,c,\epsilon}(\xi) \right| 
\leq \frac{x^{1/2 - \text{Im}(\xi)}}{\log(x)} \left( \frac{0.31}{|\text{Re}(\xi)|^3} + 1.01 \frac{|\epsilon(\xi)|}{|\text{Re}(\xi)|} \right).
\]

From Lemma 2.5 we get

\[
\sum_{|\text{Im}(\rho)| > c/\epsilon}^* \frac{1}{|\text{Im}(\rho)|^3} \leq \frac{1}{2\pi} \left( \frac{c}{\epsilon} \right)^2 \log \frac{c}{2\pi\epsilon} + 0.0001 \left( \frac{c}{\epsilon} \right)^2 \leq 0.16 \left( \frac{c}{\epsilon} \right)^2 \log \frac{c}{\epsilon}.
\]

From this, together with Lemma 2.4 and (2.19), the result follows if we take into account that 0.16 · 0.31 ≤ 0.05 and 0.65 · 1.01 ≤ 0.66.

\[\square\]

**Theorem 2.8** (Method II). Let \( x \geq 10^{10}, \epsilon \leq 10^{-5}, c \geq 10 \) and \( \delta \leq 0.1 \). Then we have

\[
(\text{2.20}) \quad \sum_{|\text{Im}(\rho)| > c/\epsilon}^* \left( \text{Ei}(\rho \log x) + \hat{\Psi}_{x,c,\epsilon,\delta} \left( \frac{e}{i} - \frac{1}{2i} \right) \right)
= \Theta \left( \frac{x^{h}}{\log x} \log \left( \frac{c}{\epsilon} \right) \left( 0.016 \left( \frac{c}{\epsilon} \right)^3 + 0.66 \left( 1 + \frac{2}{c\delta^2} \right) \log(3c) e^{(\sqrt{\pi}/4-1)c} \right) \right),
\]

where \( h = \frac{1}{2} \) if the Riemann hypothesis is assumed and \( h = 1 \) otherwise.

**Proof.** We recall that

\[
(\text{2.21}) \quad \hat{\Psi}_{x,c,\epsilon,\delta}(\xi) = \frac{x^{1/2 + i\xi}}{\log x} \left( -\frac{i}{\xi} + \frac{1}{\xi^3} \left( \frac{1}{2} - \frac{1}{\log x} \right) + \frac{i}{\xi^3} \left( \frac{1}{4} - \frac{1}{\log x} + \frac{2}{(\log x)^3} \right) \right) \left( a_\delta + b_\delta \cos(\epsilon \delta \xi) \right) l_c(\epsilon \xi) - 1),
\]

where \( a_\delta \) and \( b_\delta \) were given by

\[
(b_\delta = -\frac{\epsilon''(0)}{\delta^2} \in (0,(c\delta^2)^{-1})
\]

and \( a_\delta = 1 - b_\delta \).

Again, by a straightforward computation, we find

\[
(\text{2.22}) \quad \frac{1}{(\frac{1}{2} + i\xi)^l} + \frac{1}{((\frac{1}{2} + i\xi)^l)^2} + \frac{2}{((\frac{1}{2} + i\xi)^l)^3}
- \frac{1}{l} \left( -\frac{i}{\xi} + \frac{1}{\xi^2} \left( \frac{1}{2} - \frac{1}{l} \right) + \frac{i}{\xi^3} \left( \frac{1}{4} - \frac{1}{\log x} + \frac{2}{(\log x)^3} \right) \right)
= \frac{8 - l^2(i - 2\xi)^2 + 48i\xi - 96\xi^2 + 4l(-1 - 5i\xi + 6\xi^2)}{4l^3(i - 2\xi)^3\xi^3}.
\]

Using the same bounds as in the proof of Theorem 2.7, we get

\[
(\text{2.23}) \quad \left| \frac{8 - l^2(i - 2\xi)^2 + 48i\xi - 96\xi^2 + 4l(-1 - 5i\xi + 6\xi^2)}{4l^3(i - 2\xi)^3\xi^3} \right| \leq \frac{0.139}{l|\text{Re}(\xi)|^4}.
\]
In a similar way, we get

\begin{equation}
\text{Ei}((1/2+i\xi)l) - \text{sgn}(\text{Re}(\xi))\pi i = x^{1/2+i\xi} \left( \sum_{n=1}^{3} \frac{(n-1)!}{((1/2+i\xi)l)^{n}} + \Theta \left( \frac{0.001}{l|\text{Re}(\xi)|^{4}} \right) \right)
\end{equation}

from Lemma 2.6. Since $|a_{d} + b_{d} \cos(\varepsilon\delta\xi)| \leq 1 + 2.001/(c\delta^{2})$, we obtain the inequality

\begin{equation}
\left| \frac{1}{1} \left( -\frac{i}{\xi} + \frac{1}{\xi^{2}} \left( \frac{1}{2} - 1 \right) + \frac{i}{\xi} \left( \frac{1}{4} - \frac{1}{l^{2}} + \frac{2}{l^{2}} \right) \right) (a_{d} + b_{d} \cos(\delta\xi))l_{c}(\varepsilon\xi) \right| 
\leq 1.01 \left( 1 + \frac{1}{c\delta^{2}} \right) \frac{|l_{c}(\varepsilon\xi)|}{l|\text{Re}^{(\xi)}|}.
\end{equation}

This in combination with (2.23) and (2.24) gives

\begin{equation}
\left| \text{Ei}((1/2 + i\xi) \log x) - \text{sgn}(\text{Re}(\xi))\pi i + \hat{\Psi}_{x,c,\varepsilon}(\xi) \right|
\leq \frac{x^{1/2 - \text{Im}(\xi)}}{\log(x)} \left( \frac{0.14}{|\text{Re}(\xi)|^{3}} + 1.01 \left( 1 + \frac{1}{c\delta^{2}} \right) \frac{|l_{c}(\varepsilon\xi)|}{|\text{Re}(\xi)|} \right).
\end{equation}

Again, from Lemma 2.5 we get

\begin{equation}
\sum_{|\text{Im}(\rho)| > c/\varepsilon}^{*} \frac{1}{|\text{Im}(\rho)|^{4}} \leq 0.11 \left( \frac{\varepsilon}{c} \right)^{3} \log \frac{c}{\varepsilon}.
\end{equation}

From this, together with Lemma 2.4 and (2.26), the result follows if we take into account that $0.11 \cdot 0.14 \leq 0.016$ and $0.65 \cdot 1.01 \leq 0.66$. \hfill \Box

3. Computation of the zeros of $\zeta(s)$

We computed all zeros of $\zeta(s)$ with imaginary part in $(0, 10^{11}]$ mostly following the methods outlined in [Odl92] but using a relatively straightforward FFT method instead of the Odlyzko-Schönhage algorithm [OS88], which is normally used for fast multiple evaluations of the zeta function.

Let

\[ F(x) = \sum_{j=1}^{N} a_{j} e^{i\gamma_{j}x} \]

and suppose we want to evaluate $F(x)$ for all $x \in [-T, T] \cap \mathbb{Z}$. Let $R = 2\pi$ and

\[ \gamma_{j} = \frac{2\pi n_{j}}{R} + \delta_{j} \]

for some $\delta_{j}$ satisfying

\[ |\delta_{j}| \leq \frac{\pi}{R} \]

Next, we choose a polynomial

\[ P(t) = b_{0} + \cdots + b_{n}t^{n} \]
approximating the function \( f(t) = \exp(it) \) for \( t \in [-\pi T/R, \pi T/R] \). Then we have

\[
F(x) = \sum_{j=1}^{N} a_j e^{2\pi i n_j x/R} e^{ix\delta_j},
\]

\[
= \sum_{j=1}^{N} a_j e^{2\pi i n_j x/R} P(x\delta_j) + \Theta\left(\|f - P\|_{C^0([-\pi T/R, \pi T/R])} \sum_{j=1}^{N} |a_j|\right).
\]

We now define

\[
f_l(k) = \sum_{n_j \equiv k \mod R} a_j \delta_j \quad \text{and} \quad \hat{f}_l(x) = \sum_{k=1}^{R} f_l(k) e^{2\pi ikx/R}.
\]

Then we have

\[
\sum_{j=1}^{N} a_j e^{2\pi i n_j x/R} P(x\delta_j) = \sum_{l=1}^{n} b_l x^l \sum_{j=1}^{N} a_j \delta_j e^{2\pi i n_j x/R}
\]

\[
= \sum_{l=1}^{n} b_l \hat{f}_l(x) x^l.
\]

The function \( \hat{f}_l \) can be evaluated at all integral points in \( O(R \log(R)) \) arithmetic operations using the Fast Fourier Transform. For the polynomial \( P \) we chose the polynomials that arise from Chebyshev approximation of the function \( f \). If we take for example \( N \approx T \approx R \), it is possible, for each fixed \( \alpha > 0 \) and all \( |a_j| \) bounded by 1, to choose \( n = O(\log N) \) for a desired precision of \( N^{-\alpha} \), allowing us to evaluate \( F \) at \( N \) points in \( O(N \log(N)^2) \) compared to \( O(N^2) \) for direct evaluation. While the order of the run time of this method is the same as the one of the Odlyzko-Schönhage algorithm, our tests suggest that for the requirements of our implementation the implied constant is significantly lower.

4. Calculation of the sum over prime powers

4.1. The Fourier transform of \( l_c \). Let

\[
I_k(t) = \sum_{n=0}^{\infty} \frac{1}{n!(n+k)!} \left(\frac{t}{2}\right)^{2n+k}
\]

denote the \( k \)-th modified Bessel function of the first kind. Then we have the following:

**Proposition 4.1.** Let

\[
l_c(t) = \begin{cases} \frac{c}{2 \sinh c} I_0(c\sqrt{1-t^2}) & |t| < 1 \\ \frac{c}{4 \sinh c} & |t| = 1 \\ 0 & |t| > 1. \end{cases}
\]

Then we have \( (F^{-1}l_c)(t) = l_c(t) \).

**Proof.** Since \( l_c \) is an even function its inverse Fourier transform is also even, and thus it is sufficient to show this identity for negative values. By (\ref{1.4}), we have

\[
F^{-1}l_c(t) = \frac{c}{2\pi \sinh c} \int_{-\infty}^{\infty} e^{-it\xi} \frac{\sin(\sqrt{\xi^2 - c^2})}{\sqrt{\xi^2 - c^2}} d\xi.
\]
Now let $\xi = z + \frac{c^2}{4z}$; then we have $d\xi = (1 - \frac{c^2}{4z^2})dz$ and $\xi^2 - c^2 = \left(z - \frac{c^2}{4z}\right)^2$. Let $C$ be the path which follows the real line from $-\infty$ to $-\frac{\xi}{2}$, then crosses the upper half-plane along the circle with center $0$ and radius $\frac{\xi}{2}$ and then follows the real line again from $\frac{\xi}{2}$ to $\infty$. Then we get

\[
(4.3) \quad \int_{-\infty}^{\infty} e^{-it\xi} \frac{\sin(\sqrt{\xi^2 - c^2})}{\sqrt{\xi^2 - c^2}} d\xi = \int_{C} e^{-it(z + \frac{c^2}{4z})} \frac{\sin(z - \frac{c^2}{4z})}{z} dz
\]

\[
= \frac{1}{2i} \int_{C} e^{iz(1-t) - i\frac{c^2}{4z}(1+t)} \frac{dz}{z} - \frac{1}{2i} \int_{C} e^{-iz(1+t) + i\frac{c^2}{4z}(1-t)} \frac{dz}{z}.
\]

For $|z| \to \infty$ we have

\[
\left|\frac{e^{iz(1-t) - i\frac{c^2}{4z}(1+t)}}{z}\right| \ll \frac{e^{\operatorname{Re}(z)(t-1)}}{|z|}
\]

and

\[
\left|\frac{e^{-iz(1+t) + i\frac{c^2}{4z}(1-t)}}{z}\right| \ll \frac{e^{\operatorname{Im}(z)(1+t)}}{|z|}.
\]

Consequently, both integrals on the right hand side of (4.3) vanish for $t < -1$ as the path of integration is shifted towards $i\infty$, and thus $F^{-1}l_c(t)$ vanishes for $|t| > 1$. For $t \in (-1, 0)$ the first integral in (4.3) still vanishes as the path of integration is shifted towards $i\infty$, and the same holds for the second integral as the path of integration is shifted towards $-i\infty$. Therefore, we find that

\[
F^{-1}(l_c)(t) = \frac{c}{2\sinh c} \operatorname{Res}_{z=0} \frac{e^{-iz(1+t) + i\frac{c^2}{4z}(1-t)}}{z}
\]

holds for such $t$. The residue equals the constant term in the Laurent series expansion for $e^{-iz(1+t) + i\frac{c^2}{4z}(1-t)}$ at $0$. We have

\[
e^{-iz(1+t) + i\frac{c^2}{4z}(1-t)} = \sum_{n=0}^{\infty} \frac{i^n}{n!} \left(-z(1+t) + \frac{c^2}{4z}(1-t)\right)^n
\]

\[
= \sum_{n=0}^{\infty} \frac{i^n}{n!} \sum_{k=0}^{n} \binom{n}{k} (-1)^k \left(\frac{c^2}{4z}(1-t)\right)^{n-k} (z(1+t))^k
\]

and therefore get

\[
\operatorname{Res}_{z=0} \frac{e^{-iz(1+t) + i\frac{c^2}{4z}(1-t)}}{z} = \sum_{n=0}^{\infty} \frac{1}{(n!)^2} \left(\frac{c\sqrt{1-t^2}}{2}\right)^{2n} = I_0(c\sqrt{1-t^2}).
\]

The value at $t = \pm 1$ is obtained since $F^{-1}l_c$ is normalized by a well-known fact from Fourier analysis. \hfill \Box

4.2. The Fourier transforms of $\Phi_{x,c,\varepsilon}$ and $\Psi_{x,c,\varepsilon,\delta}$. We define the auxiliary functions

\[
n_j(\xi) = \frac{l_c(\xi) - 1}{(i\xi)^j} \quad \text{and} \quad m_j(\xi) = \frac{g_{c,\delta}(\xi) - 1}{(i\xi)^j}.
\]

For $j \in \{0, 1, 2\}$ we denote by $\nu_j$ the inverse Fourier transform of $n_j$, and for $j \in \{0, 1, 2, 3\}$ we denote by $\mu_j$ the inverse Fourier transform of $m_j$. Then we get the following proposition.
Proposition 4.2. The inverse Fourier transforms of $\hat{\Phi}_{x,c,\varepsilon}$ and $\hat{\Psi}_{x,c,\varepsilon,\delta}$ are given by

$$\Phi_{x,c,\varepsilon}(t) = \sqrt{x} \log x \left( \nu_1 \left( \frac{t - \log x}{\varepsilon} \right) - \varepsilon \nu_2 \left( \frac{t - \log x}{\varepsilon} \right) \left( \frac{1}{2} - \frac{1}{\log x} \right) \right)$$

and

$$\Psi_{x,c,\varepsilon,\delta}(t) = \sqrt{x} \log x \left( \mu_1 \left( \frac{t - \log x}{\varepsilon} \right) - \varepsilon \mu_2 \left( \frac{t - \log x}{\varepsilon} \right) \left( \frac{1}{2} - \frac{1}{\log x} \right) \right.$$  

$$+ \varepsilon^2 \mu_3 \left( \frac{t - \log x}{\varepsilon} \right) \left( \frac{1}{4} - \frac{1}{\log x} + \frac{2}{(\log x)^2} \right) \right).$$

Consequently, we have

$$\supp(\Phi_{x,c,\varepsilon}) \subset B_{\varepsilon}(\log(x)) \quad \text{and} \quad \supp(\Psi_{x,c,\varepsilon,\delta}) \subset B_{(1+\delta)\varepsilon}(\log(x)).$$

Proof. This follows immediately if we take into account that

$$F^{-1} \left( f(\alpha \cdot) \right) = \frac{1}{|\alpha|} \left( F^{-1} f \right) \left( \frac{\cdot}{\alpha} \right)$$

when $\alpha \neq 0$ and

$$F^{-1} \left( \exp(i\beta \cdot f(\cdot)) \right) = (F^{-1} f)(\cdot - \beta).$$

Next, we give a description of how to compute $\nu_j$ and $\mu_j$. Let $L_0 = \iota_c$ and for $j \geq 0$ let

$$L_{j+1}(t) = \int_{-\infty}^{t} L_j(\tau) \, d\tau,$$

i.e. those antiderivatives of $\iota_c$ which vanish for $t < -1$. The functions $\mu_j$ and $\nu_j$ satisfy the relations

$$\nu_j' = \nu_{j-1} \quad \text{and} \quad \mu_j' = \mu_{j-1},$$

which in the case $j = 1$ have to be understood as identities of distributions. Furthermore, we have

$$\nu_0 = L_0 - \delta_0 \quad \text{and} \quad \mu_0 = a_\delta L_0 + \frac{b_\delta}{2} (L_0(\cdot - \delta) + L_0(\cdot + \delta)) - \delta_0,$$

where $\delta_0$ denotes the Dirac distribution defined by $\delta_0(f) = f(0)$. We define

$$g_1(t) = \chi_{[0,\infty)}(t), \quad g_2(t) = \frac{1}{2}(|t| + t), \quad \text{and} \quad g_3(t) = \frac{t}{4}(|t| + t),$$

i.e. those antiderivatives of $\delta_0$ which vanish for $t < 0$, and get

$$\nu_j = L_j - g_j \quad \text{and} \quad \mu_j = a_\delta L_j + \frac{b_\delta}{2} (L_j(\cdot - \delta) + L_j(\cdot + \delta)) - g_j.$$

Now let

$$\lambda_k = -\frac{(-2)^{k+1}}{k! \sinh c} I_k(c).$$

Then, using (4.1), it is easy to see that

$$\frac{c}{2 \sinh(c)} I_0(c\sqrt{1 - t^2}) = \sum_{k=0}^{\infty} \lambda_k t^{2k}.$$
Then we get a power series representation of \( L_j \) in \((-1, 1)\) by integrating the sum in (4.3). In order to determine the constants of integration we note that

\[
L_{2j+1}(0) = \int_{-1}^{0} L_{2j}(t) \, dt = \frac{1}{(2j)!} \int_{-1}^{0} t^{2j} I_c(t) \, dt = \frac{1}{2(2j)!} \int_{-1}^{0} t^{2j} I_c(t) \, dt = \frac{1}{2(2j)!} \left( \frac{1}{2} I_c^{(2j)}(0) \right),
\]

which follows by partial integration. So we get

\[
L_1(0) = \frac{1}{2} \quad \text{and} \quad L_3(0) = \frac{1}{2} \left( \frac{1}{c \tanh(c)} - \frac{1}{c^2} \right).
\]

For \( L_2(0) \) we have

\[
L_2(0) = \int_{-1}^{0} L_1(t) \, dt = [L_0(t)]_{-1}^{0} - \int_{-1}^{0} t L_0(t) \, dt = -\frac{c}{2 \sinh c} \int_{-1}^{0} t I_0(c \sqrt{1-t^2}) \, dt.
\]

The substitution \( \tau = c \sqrt{1-t^2} \) then yields

\[
\int_{-1}^{0} t I_0(c \sqrt{1-t^2}) \, dt = \frac{1}{c} \int_{-c}^{0} \tau I_0(\tau) \, d\tau = -\frac{I_1(c)}{c},
\]

where we used \((t I_1(t))' = t I_0(t)\). So we get

\[
L_2(0) = \frac{I_1(c)}{2 \sinh(c)}.
\]

4.3. Computation of \( w_f(\Phi_{x,c,\varepsilon}) \). Let \( N_0 = \lfloor \exp(-\varepsilon)x \rfloor \), \( N_1 = \lceil \exp(\varepsilon)x \rceil \). We define

\[
\Omega_1(t) = \frac{\log(t)}{\sqrt{t}} \Phi_{x,c,\varepsilon}(\log(t)).
\]

Then we have

\[
w_f(\Phi_{x,c,\varepsilon}) = -\sum_{N_0 \leq p^m \leq N_1} \frac{\Omega_1(p^m)}{m}.
\]

The computation of \( w_f(\Phi_{x,c,\varepsilon}) \) splits into two tasks: generating all prime powers \( p^k \in I = [N_0, N_1] \cap \mathbb{Z} \) and evaluating \( \Omega_1(p^k) \).

We already mentioned in the previous subsection that \( \Phi_{x,c,\varepsilon} \) can be evaluated by power series expansion. The coefficients can be estimated with Luke’s inequality

\[
1 < k! \left( \frac{2}{t} \right)^k I_k(t) < \cosh t
\]

(see [Luk72]), which gives

\[
|\lambda_k| \leq \left( \frac{c}{2} \right) \frac{2k}{(k!)^2},
\]

for \( c \geq 1 \). For \( k > ec \) the right hand side is smaller than \( c 4^{-k} \) so that we have

\[
\sum_{l=k}^{\infty} \lambda_l | \leq c \sum_{l=k}^{\infty} 4^{-l} = \frac{3c}{4^k},
\]

for such \( k \). If we take \( k = \lfloor ec \rfloor \), then the right hand side of (4.8) gets smaller than \( 2^{-76} \) for \( c \geq 15 \), which is sufficiently small for computations with \( x \leq 10^{30} \). So
let $K = \lceil ce \rceil$. Then $I_K(c), I_{K-1}(c), I_0(c)$ and $I_1(c)$ are evaluated using the power series expansions (4.1), and for the remaining $I_k(c)$ the recurrence formula

$$I_{k-1}(c) - I_{k+1}(c) = \frac{2k I_k(c)}{c}$$

is used. For $c \leq 50$ the functions $I_k(c)$ already take values up to $2^{70}$, and the recurrence formula amplifies round-off errors for $2k > c$. Therefore, this part of the program uses mpfr-variables with 256 bit precision.

For $m \geq 2$ the prime powers $p^m$ are generated with the Eratosthenes sieve, and $\Omega_1(p^m)$ is evaluated directly using the power series expansion. To lower the expense of the evaluation of $\Omega_1(p)$, $I$ is segmented into subintervals

$$I_k = [x + kh, x + (k + 1)h) \cap I,$$

on each of which polynomial interpolation is used. For values of $x$ larger than $10^{22}$, generating the prime numbers in $I_k$ by sieving with all prime numbers $\leq \sqrt{N_1}$ gets inefficient because of the large space requirement. The space requirement can be lowered in the following way:

For any interval $I = [N_0, N_1]$ we define

$$S_{0,P}(I) = \{ n \in I \mid p | n \Rightarrow p > P \}$$

and

$$S_{1,P}(I) = \{ n \in I \mid n = pq, p, q > P \}.$$

We refer to $S_{0,P}(I)$ as the survivors of sieve 0 and to $S_{1,P}(I)$ as the survivors of sieve 1. Then, for $N_1^{1/3} \leq P \leq \sqrt{N_0}$, we have

$$S_{0,P}(I) \setminus S_{1,P}(I) = \{ n \in I \mid n \text{ is prime} \}.$$

This gives an advantage in space requirement since we can generate $S_{1,P}(I)$ by generating $S_{0,P}(\lfloor \frac{N_0}{P}, \lfloor \frac{N_1}{P} \rfloor \rfloor)$ for every prime $p \in (P, N_1/P]$. This way, we only have to keep $O(P + N_1/P)$ prime numbers in memory at a time. This reduces the space requirement for generating all primes in $I$ to $O(x^{1/3})$ but at the expense of increasing the run time to $O(x^{2/3+\eta})$. This problem could be solved by using a dissected Atkin-Bernstein sieve as described in [Gal04].

For sieve 0, we take $h = 2^{20}$, which is chosen to make each $I_k$ fit in the L1 cache, and a polynomial degree of 2. The $I_k$ are sieved one by one, and the data

$$s_{k,j} = \sum_{n \in S_{0,P}(I_k)} \left( n - x - kh - \frac{h}{2} \right)^j$$

for $j \leq 2$ is computed, stored to hard disk and later retrieved to approximate the accumulated values of $\Omega_1$ at the survivors of sieve 0.

The situation for sieve 1 differs from the situation for sieve 0 insofar as all $I_k$ have to be kept in memory at the same time. Therefore, sieve 1 is applied with a larger value for $h$ and a higher degree in the approximation formula.

To estimate the interpolation error, we have to compute bounds for the derivatives of $\Omega_1$. Since doing this by hand gets cumbersome for higher degrees, we wrote a little program to perform this task.
First we give a bound for the derivatives of \( \nu \). It is easy to see from the power series expansion (4.1) that

\[
\frac{d}{dy} I_k(y) = \frac{I_{k+1}(y)}{y^k}.
\]

So if we take \( y = c \sqrt{1 - t^2} \), we get

\[
\frac{d}{dt} I_k(y) = -tc^2 \frac{I_{k+1}(y)}{y^{k+1}}.
\]

This way, we see that

\[
\frac{d^k}{dt^k} I_0(y) = \sum_{k/2 \leq l \leq k} \beta_l^{(k)} c^{2l} t^{2l-k} \frac{I_l(y)}{y^l},
\]

where \( \beta_l^{(k)} \) is inductively defined by

\[
\beta_0^{(k)} = \delta_{l,0} \quad \text{and} \quad \beta_l^{(k+1)} = (2l - k) \beta_l^{(k)} - \beta_{l-1}^{(k)}.
\]

So taking into account that \( I_k(t)/t^k \) is monotonously increasing for \( t > 0 \) we get

\[
\left| \frac{d^k}{dt^k} I_0(y) \right| \leq \sum_{k/2 \leq l \leq k} \left| \beta_l^{(k)} \right| c^l I_l(c)
\]

for \( y \leq c \). Now let \( z = \log(t/x)/\varepsilon \). Then we have

\[
\frac{d^k}{dt^k} f(z) = \frac{1}{t^k} \sum_{l=1}^{k} \frac{u_l^{(k)}}{\varepsilon^l} f^{(l)}(z)
\]

for \( k \geq 1 \), where \( u_l^{(k)} \) is inductively defined by

\[
u_1^{(1)} = \delta_{1,l} \quad \text{and} \quad u_l^{(k+1)} = u_l^{(k)} - ku_l^{(k)}.
\]

Furthermore, we have

\[
\frac{d^k}{dt^k} \log t = \frac{k!}{2\pi i} \int_{|\xi - t| = \frac{3}{2}} \frac{\log \xi}{\sqrt{\xi} (\xi - t)^{k+1}} d\xi = \Theta \left( k! 2^{k+1/2} \frac{\log t + 2 - \frac{t}{2}}{t^{k+1/2}} \right)
\]

for \( t > 0 \), where we used

\[
|\log \xi| \leq \log(\frac{3}{2} t) + \pi/2 < \log t + 2 \quad \text{and} \quad \left| \sqrt{\xi} (\xi - t)^{k+1} \right| \geq \frac{t^{k+3/2}}{2^{k+3/2}}
\]

The derivatives of \( \Omega_1 \) can thus be estimated using the Leibnitz formula. For the computation of \( \pi(10^{24}) \), we took \( c = 30, \varepsilon = 3 \times 10^{-10} \). For the interpolation of the sieve 0 data (where we took \( h = 2^{20} \) and a polynomial degree of 2) the aforementioned procedure shows that the interpolation error is \( \leq 7 \times 10^{-25} \) and the accumulated error is \( \leq 4 \times 10^{-10} \). This took an estimated 1,337 CPU hours on the 2.66 GHz Xeon CPU mentioned at the end of the introduction. The estimate is based upon extrapolation (which is quite accurate for the sieve 0 program), since in reality the jobs were split among several different CPUs.

For sieve 1 we chose \( h = 2^{36} \) and a polynomial degree of 6, in which case the interpolation error for a single evaluation is \( \leq 5 \times 10^{-22} \) and the accumulated error is \( \leq 3 \times 10^{-7} \). Total time of sieve 1 jobs was 221 CPU hours on several Intel Xeon CPUs in the 2.66–3.0 GHz range.
4.4. Computation of $w_\infty(\Phi_{x,c,\varepsilon})$. Since the function $\nu_1(t)$ is odd and vanishes for $|t| > 1$, it is bounded by
\[
\int_{-1}^{0} \nu_c(t) \, dt = \frac{1}{2} \int_{-\infty}^{\infty} \nu_c(t) \, dt = \frac{l_c(0)}{2} = \frac{1}{2}.
\]
Consequently $\nu_2$, as an antiderivative for $\nu_1$, which also vanishes for $|t| > 1$, is also bounded by $1/2$ in absolute value.

Therefore, we have
\[
|\Phi_{x,c,\varepsilon}(t)| \leq \sqrt{x} \log x,
\]
for $x > e$ and $\varepsilon < 1$, and so the inequality
\[
|w_\infty(\Phi_{x,c,\varepsilon})| \leq \frac{\sqrt{x}}{\log x} \int_{\log x - \varepsilon}^{\log x + \varepsilon} \frac{e^{-\frac{1}{2}t}}{1 - e^{-2t}} dt \leq 10 \frac{\varepsilon}{\log x}
\]
holds for $x \geq 100$. Consequently, it is not necessary to compute $w_\infty(\Phi_{x,c,\varepsilon})$ in order to compute $\pi(x)$, but it can be very useful for debugging purposes to know its value. Also, for the first implementation one would like to use large values of $\varepsilon$ to find possible bugs. To this end, we calculated the integral with the numerical integration routines of the GNU scientific library. The same goes for the integral
\[
\int_{x}^{\infty} \frac{dt}{t \log(t^2 - 1)},
\]
which occurs in the Riemann explicit formula and which is bounded by $\frac{1}{x^2 \log x}$.

5. Computation of $\pi(x)$ with $\Psi_{x,c,\varepsilon,\delta}$

We give a short description of the second method that computes $\pi(x)$ by applying Weil’s explicit formula to $\Psi_{x,c,\varepsilon,\delta}$. The main difference from the first method is that the function $\Psi_{x,c,\varepsilon,\delta}$ has salient points at $\exp(\pm \varepsilon)x$ which have to be considered in the estimation of the interpolation error.

5.1. Computation of $w_f(\Psi_{x,c,\varepsilon,\delta})$. Let
\[
M_0 = \lfloor \exp(-(1 + \delta)\varepsilon)x \rfloor, \quad M_1 = \lfloor \exp((1 + \delta)\varepsilon)x \rfloor
\]
and
\[
\Omega_2(t) = \frac{\log t}{\sqrt{t}} \Psi_{x,c,\varepsilon,\delta}(t).
\]
The derivatives of $\Omega_2$ can be estimated using the same procedure as in the previous section, but the derivatives now have discontinuities at $N_0$ and $N_1$, which have to be considered separately. This can be done as follows.

Let $f$ be a function, which is continuous in $[0, h]$ and smooth in $(0, h) \setminus \{s\}$ with bounded derivatives, and let
\[
A_n = \lim_{\varepsilon \to 0} (f^{(n)}(t_s + \varepsilon) - f^{(n)}(t_s - \varepsilon)).
\]
We define $g_n(t) = \chi_{(0,\infty)}(t)t^n/n!$. Then the function

$$f_d(t) = f(t) - \sum_{n=1}^{d} A_n g_n(t - t_s)$$

is an element of $C^d((0,1))$, and we have

$$\|f_d\|_{C^0} \leq \|f_d\|_{\infty} + |A_d|.$$ 

Now let $I_d$ be an interpolation operator of degree $d$, i.e. an operator which sends $f$ to the uniquely determined polynomial of degree $\leq d$ which interpolates $f$ in $d + 1$ distinct points of $[0,\h]$. Then we have

$$\|f - I_d(f)\|_{\infty} \leq \|f - f_{d+1}\|_{\infty} + \|f_{d+1} - I_d(f_{d+1})\|_{\infty} + \|I_d(f_{d+1}) - I_d(f)\|_{\infty}$$

$$\leq (1 + \|I_d\|)\|f - f_{d+1}\|_{\infty} + \|f_{d+1} - I_d(f_{d+1})\|_{\infty}.$$

For arbitrary interpolation points the norm of $I_d$ can be arbitrarily large, but if we restrict ourselves to Chebyshev interpolation, i.e. take the interpolation points $h \cos(\pi(j+1)/(d+1))$, it can be seen from the discrete orthogonal relation of the Chebyshev polynomials that $\|I_d\| \leq 2^{d+1}$. Furthermore, the $n$-th normalized Chebyshev polynomial is bounded by $2^{1-n}$ on $[-1,1]$, so combining this with the standard estimate for polynomial interpolation, we get

$$\|f - I_d(f)\|_{\infty} \leq h^{d+1}\frac{2}{d+1} + \|A_{d+1}\| + (2d + 2) \sum_{n=1}^{d+1} |A_n| h^n/n!.$$

5.2. Computation of $w_\infty(\Psi_{x,c,\varepsilon,\delta})$. The term $w_\infty(\Psi_{x,c,\varepsilon,\delta})$ is again very small. Under the conditions of Theorem 2.8 ($c \geq 10, \varepsilon \leq 10^{-5}, \delta \leq 0.1$), we get

$$|\Psi_{x,c,\varepsilon}(t)| \leq \frac{\sqrt{x}}{\log x} \left(2 + \frac{4}{c\delta^2}\right)$$

from (4.2) and (4.4). So, as in (4.11), we get

$$|w_\infty(\Psi_{x,c,\varepsilon,\delta})| \leq \left(10 + \frac{20}{c\delta^2}\right)\frac{\varepsilon}{\log x}$$

for $x \geq 100$.

6. Comparison to the Galway-Platt method

The question whether the Logan function used in this paper or the Gaussian function used in the Galway-Platt method is more suitable for calculating $\pi(x)$ is not easily answered. The Logan function has a sharp cutoff property, as mentioned earlier in section 1. So using the Logan function can be seen as the holomorphic analog of cutting off the sum over zeros at the point $c/\varepsilon$, neglecting the terms that are not affected by the Logan function and which impose a lower bound on the zero range. Furthermore, the use of the Logan function results in a finite correcting sum over prime powers, so one does not have to truncate this sum as in [Gal04].

A crucial question is the method’s efficiency. The choice of the Gaussian function for the Galway-Platt method appears to be motivated by the fact that both the function and its Fourier transform decay rapidly and the Gaussian function is in a certain way optimal in this respect [Gal04, section 2.5]. However, the efficiency of the method is rather ruled by the lengths of the zero range and the sieve interval, which are practically in inverse proportion for reasonable choices of parameters.
Consequently, their geometric mean is constant and gives a good measure of the methods’ efficiency. With respect to this measure, the Logan function seems to give equally good or better results:

For the Galway-Platt method the sieve interval is taken to be \([e^{-\tau}x, e^{\tau}x]\), where

\[
\tau \sim \lambda \sqrt{2 \log \left( \frac{\lambda x}{\eta} \right)}
\]

[Gal04, Theorems 3.5 and 3.10]. If one uses

\[
\alpha_T \approx 1 + \frac{\log \left( \frac{1}{\pi} \log(T) \right)}{\log(T)}
\]

in Lemma A.4 and Lemma A.5 in [Pla15], the main terms take the form

\[
\frac{1}{h \pi \log x} \log(T) \frac{e^{-T^2 \lambda^2/2}}{T^2 \lambda^2},
\]

where \(h = \frac{1}{2}\) if the Riemann hypothesis is assumed and \(h = 1\) otherwise. So calculating the sum over zeros unconditionally within a sufficient accuracy can be achieved by choosing \(T = \sqrt{2x \log(x/\eta)}\) and \(\lambda = 1/\sqrt{x}\) for a sufficiently small \(\eta > 0\).

Consequently, the geometric mean for the Galway-Platt method is asymptotically

(6.1) \(2^{3/4} \sqrt{x \log(x)}\).

For the methods in this paper the choice \(c = \log(x/\eta) + \log \log \log(x)\) and \(\varepsilon = 1/\sqrt{x}\) (and \(\delta = c^{-1/2}\) in Method II) yields a geometric mean which is asymptotically

(6.2) \(2^{1/2} \sqrt{x \log(x)}\).

The constants change if the Riemann hypothesis is assumed. In case of the Gaussian function one can save a factor \(\sqrt{2}\) in (6.1), where it suffices to assume this hypothesis for \(\Im(\rho)\) roughly up to \(\sqrt{2T}\). For the Logan function one can even save a factor \(\sqrt{2}\) in (6.2) assuming full RH. This is a benefit of the additional parameter \(c\), which provides more flexibility for controlling the size of the remainder. It can also be shown by a refined analysis of the sum over zeros that one can save a factor \((3/4)^{1/4}\) in (6.2) if the Riemann hypothesis is assumed up to \(2T/\sqrt{3}\), which is done in a subsequent paper [Büt].

The aforementioned measure may also be used for a direct comparison of the calculations of \(\pi(10^{24})\), which supports the asymptotic results. Method II could perform the calculation in [Pla15] (which already uses partial knowledge of the RH) equally efficiently, taking \(c = 61, \varepsilon = 6.1 \times 10^{-10}\) and \(\delta = 0.05\). For the conditional computation with method I the geometric mean turns out to be about 30% smaller.

On the plus side, the Galway-Platt method is less restrictive in choosing the length of the zero range, and the functions occurring in the sums over zeros and prime powers are a bit simpler.

References


W. F. Galway, Analytic Computation of the Prime-Counting Function, ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)–University of Illinois at Urbana-Champaign, 2004. MR2706687

D. H. Lehmer, On the exact number of primes less than a given limit, Illinois J. Math. 3 (1959), 381–388. MR0106883 [21 #5613]


B. Rosser, Explicit bounds for some functions of prime numbers, Amer. J. Math. 63 (1941), 211–232. MR0003018 [12,150c]
